Pip-HoGu: An Artificial Assembly with Cooperative DNA Recognition Capable of Mimicking Transcription Factor Pairs

AUTHOR(S):

Yu, Zutao; Guo, Chuanxin; Wei, Yulei; Hashiya, Kaori; Bando, Toshikazu; Sugiyama, Hiroshi

CITATION:

Yu, Zutao ...[et al]. Pip-HoGu: An Artificial Assembly with Cooperative DNA Recognition Capable of Mimicking Transcription Factor Pairs. Journal of the American Chemical Society 2018, 140: 2426-2429

ISSUE DATE:

2018-02-21

URL:

http://hdl.handle.net/2433/230818

Pip-HoGu, an artificial assembly with cooperative DNA recognition capable of mimicking transcription factor pairs

Zutao Yu†, Chuanxin Guo†, Yulei Wei†, Kaori Hashiya†, Toshikazu Bando†, Hiroshi Sugiyama*†‡

† Department of Chemistry, Graduate School of Science and ‡Institute for Integrated Cell-Material Sciences (iCeMS), Kyoto University, Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan

Supporting Information Placeholder

ABSTRACT: Cooperation between pairs of transcription factors (TFs) has been widely demonstrated to play a pivotal role in the spatiotemporal regulation of gene expression, but blocking cooperative TF pair–DNA interactions synergistically has been challenging. To achieve this, we designed programmable DNA binder pyrrole-imidazole polyamides conjugated to host–guest assemblies (**Pip-HoGu**) to mimic the cooperation between natural TF pairs. By incorporating cyclodextrin (Cyd)–adamantane (Ada), we synthesized **Ada1** (PIP1-Ada) and **Cyd1** (PIP2-Cyd), which were evaluated using T_m, EMSA, competitive, and SPR assays and molecular dynamics studies. The results consistently demonstrated that **Pip-HoGu** system formed stable noncovalent cooperative complexes, thereby meeting key criteria for mimicking a TF pair. The system also had a longer recognition sequence (two-PIP binding length plus gap distance), favorable sequence selectivity, higher binding affinity, and in particular, a flexible gap distance (0–5 bp). For example, **Ada1**–**Cyd1** showed thermal stability of 7.2 °C and a minimum free energy of interaction of –2.32 kcal·mol^{-1} with a targeting length of 14 bp. Furthermore, cell-based evaluation validated the capability of **Pip-HoGu** to exhibit potent cooperative inhibitory effects on gene expression under physiological conditions by disrupting TF pair–DNA function. In conclusion, the modular design of **Pip-HoGu** defines a general framework for mimicking naturally occurring cooperative TF pair–DNA interactions that offers a promising strategy for applications in the precise manipulation of cell fate.

Manipulating spatiotemporally variable gene expression has been the goal of generations of scientists[1,2]. In mammals, there are approximately 1000 transcription factors (TFs) that extensively regulate gene expression patterns, and 55–70% of these TFs may be functioning as cooperative TF pairs via homo-/heterodimers to ensure high binding affinity and extended recognition sequence[3,4]. Programmable molecules, e.g., nucleic acid analogues and pyrrole-imidazole polyamides (PIPs), can disrupt individual TF–DNA interactions[5-7], but cannot block interactions between collaborative TF pairs and DNA. More specifically, the gap sequences between the two binding motifs of the TF pair are not conserved and the gap distances are relatively flexible, ranging from –1–5 bp[8]. Most significantly, by switching cooperative partners, TF pairs can exert divergent biological functions. For example, Sox2/Oct4 instigates pluripotent gene activation, but Oct4/Sox17 functions as a HEX activator and Sox2/Pax6 plays a pivotal role in ocular lens development[9,10]. Accordingly, disrupting the individual binding sites of Sox2 or Oct4, rather than synergistically disrupting those of both TFs will result in complex biological outcomes. Therefore, novel strategies are needed to address these challenges to the implementation of deliberate and precise manipulation of gene expression patterns.

PIPs are the best characterized programmable DNA minor-groove binders that can compete with TF binding, with the binding rules that Py/Im recognizes C/G, Im/Py recognizes G/C, and Py/Py recognizes A/T and T/A[11]. Despite substantial progress, there is still a long way to go before these molecules can be applied clinically[7,12]. The short recognition sequences of PIPs lead to high off-target rates, but the extension of PIP length significantly impairs its cell permeability[13,14]. Critically, the fixed PIPs-binding motif (4–8 bp) without elasticity, is unsuitable for blocking flexible TF binding, especially that of TF pairs[15].

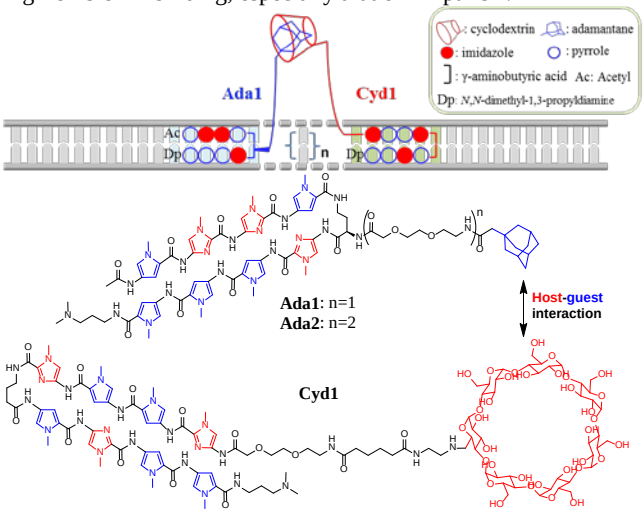


Figure 1. Overview of cooperative interactions of a TF pair targeting a sequence associated with two components of **Pip-HoGu** assembly, **Ada1** and **Cyd1**. n = gap distance. (Bottom) Chemical structures of **Ada1**, **Ada2**, and **Cyd1**.

In this context, we envisaged the integration of PIPs with a cooperative system to mimic the homo- or heterodimer binding systems of TF pairs. There are several classic noncovalent cooperative systems, including nucleic acid analogues, metal ion–ligand, and host–guest systems (e.g., cyclodextrin (Cyd), cucurbit[n]uril, and carcerands with guests)[16-21]. Among these cooperative systems, Cyd–adamantane (Ada) has been extensively studied as an exemplary host–guest system both in vitro and in cells[22,23]. By replacing a leucine-zipper dimerization domain with Cyd/Ada, Morii and colleagues designed an artificial system in which the cooperative Cyd–Ada interaction highly stabilized the interaction

of DNA with the DNA-binding domains of GCN4 homodimer[24,25]. The work of Mascarenas and colleagues of Cyd–Ada assisted DNA-binding peptide–distamycin derivatives represented a step forward in the development of smaller, selective, and ligand-responsive systems[26]. Accordingly, the design of programmable DNA binder/host–guest scaffold for mimicking cooperative TF-pair systems and especially their cell-based applications is highly attractive.

To achieve this, we designed PIPs conjugated to a host–guest Cyd–Ada scaffold, i.e., **Pip-HoGu**. We first evaluated them in vitro using the DNA-binding sequences of the Tax/CREB heterodimer, which functions by cooperative recruitment of p300 that is essential for HTLV-1 virus amplification[27,28] **Ada1** (PIP1-Ada) consists of a PIP moiety to target the Tax binding site (5′-WWGGCW-3′) conjugated to Ada via a mini-PEG linker (Figure 1)[24]. Host conjugate **Cyd1** (PIP2-Cyd) contains a Cyd moiety and a CREB-competitive-binding PIP (5′-WGWCGW-3′). We generated series of positive- and negative-binding sequences, and the difference in binding originates from the relative positions of the **Ada1** and **Cyd1** binding sites (Figure 2A). In positive-binding sequences, Cyd–Ada covers only a short distance (equal to the gap distance) (Figure S1). In contrast, in the negative-binding mode, Ada must bridge two PIP-binding sites plus the gap distance, making it impossible for Ada to interact with Cyd.

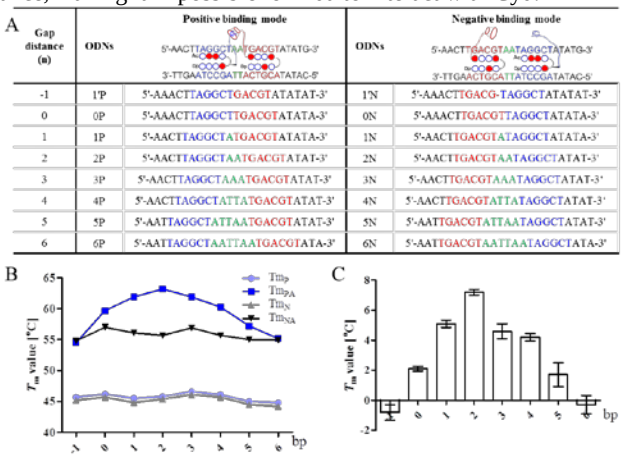


Figure 2. T_m assay illustrating the cooperativity of **Pip-HoGu**. (A) The DNA oligomers (ODNs) used in the T_m assay, including positive (ODN1′P–ODN6P) and negative (ODN1′N–ODN6N) binding sequences. The gap distance (green) is the number of bp between the binding sites of **Ada1** (blue) and **Cyd1** (red). The chart only shows the forward DNA strand. (B) T_m profiles of positive ODNs (T_{mP}, light blue), negative ODNs (T_{mN}, gray), positive ODNs/**Ada1–Cyd1** (T_{mPA}, blue), and negative ODNs/**Ada1–Cyd1** (T_{mNA}, black). (C) $\Delta\Delta T_m$ profiles of cooperativity of **Ada1–Cyd1** assemblies. $\Delta T_m = T_m$ (ODNs/PIPs) – T_m (ODNs); $\Delta\Delta T_m = \Delta T_{mP} - \Delta T_{mN}$. Error bars indicate the standard deviation of three replicates.

A thermal stabilization assay (T_m assay) was performed to evaluate the cooperative binding potency and how it was influenced by the gap distance[29]. In the positive-binding mode, the overall thermal stability of **Ada1**–**Cyd1** had a ΔT_{mP} value of 9–15 °C ($\Delta T_{mP} = T_{mP} - T_m$) in a gap-distance-dependent manner (Figure 2B, Table S1). In negative-binding mode, however, there were no gap-distance-dependent effects on the thermal stability of **Ada1**–**Cyd1** with ΔT_{mN} values around 9–10 °C **(**Table S1). In the control experiment with mixture of PIP1 and PIP2, there was no significant difference of thermal stability between positive- and negative-binding sequences (Figure S2). Therefore, the discrepancy of thermal stabilization effect between positive- and negative-binding modes should mainly attributable to the cooperative interaction of the Cyd–Ada complexes[30,31]. The results showed that positive-binding ODNs with 0–5 bp gap distances displayed cooperative binding function, and no cooperative effect was observed with gap distances ≥6 bp (Figure 2C). These results highlight the gap-distance dependency of cooperative binding energies. ODN2P with a 2-bp gap distance demonstrated the highest level of cooperation ($\Delta\Delta T_m$ = 7.2 °C). Of note, a 1-bp mismatch T_m assay showed that **Pip-HoGu** exhibited high sequence specificity with a $\Delta\Delta T_m$ of 8.3 °C (Table S2)[32].

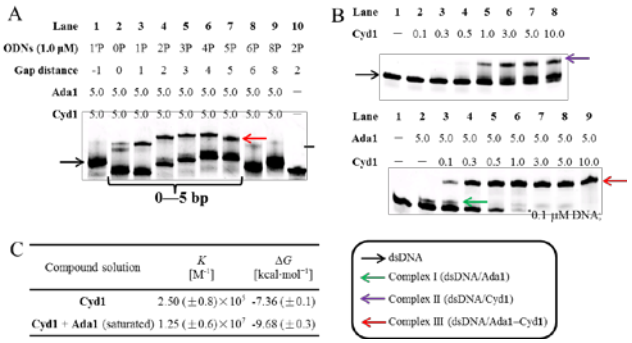


Figure 3. EMSA illustrating the cooperativity of **Pip-HoGu**. (A) The gel-shift behavior of all the positive-binding sequences with **Ada1–Cyd1**. Concentrations are shown in figure. (B) Quantitative EMSA of ODN2P with **Cyd1** at various concentrations (top) and **Cyd1** supplemented with saturated **Ada1** (bottom). ODNs concentration: 0.1 μM. (C) Equilibrium association constants and free energies for ODN2P with **Ada1–Cyd1**.

Parallel to the T_m assay, electrophoretic mobility shift assays (EMSAs) were conducted to visualize band-shift behavior upon formation of stable complexes[33]. The band upshifts for **Ada1**, **Cyd1**, and **Ada1–Cyd1** with ODN2P can be clearly distinguished (Figure S3). Next, fixed concentrations of **Ada1**–**Cyd1** were allowed to equilibrate with all ODNs. In agreement with the results of the T_m assay, the appearance of an upshifted band showed that ODNs with positive-binding mode and 0–5 bp gap distances display cooperative binding (Figure 3A). **Ada1–Cyd1** shows substantially weaker band shift with ODN0P, for which steric hindrance might be partially responsible. In comparison, no complexes were observed for DNA sequences of –1 bp and >6 bp. Moreover, no negative-binding mode DNA sequences could form upshifted band indicative of cooperative complexes, and there was also no upshifted band of negative ODNs with **Ada1** or **Cyd1** individually, suggesting cooperative complex mediated sequence selectivity (Figure S4)[3]. In addition, competitive EMSA assays showed that cooperation was weakened in the presence of a guest competitor (Figure S5)[34].

Quantitative EMSAs were performed to analyze the magnitude of cooperativity and the equilibrium association constant was determined by fitting to the Langmuir binding isotherm[35]. The increase in the upshifted band for ODN2P at various concentrations of **Cyd1** alone and in the presence of **Ada1** at excess concentration demonstrates the cooperative effect (Figure 3B)[36]. Specifically, the data generated an equilibrium association constant of 2.50×10^5 M^{-1} (K_1) for **Cyd1** alone, and promisingly increased to 1.25×10^7 M^{-1} ($K_{1,2}$) in the presence of **Ada1** (Figure 3C). Using the equation for the free energy of binding, the free energies of binding for **Cyd1** alone and in the presence of **Ada1** were –7.36 and –9.68 kcal·mol^{-1}, respectively, giving a minimum free energy of interaction ($G_{2\text{-}1} - G_2$) of –2.32 kcal·mol^{-1}[37]. Therefore, **Pip-HoGu** has superior cooperation-stabilization effects to the previously reported 8-bp DNA duplex (–2.2 kcal·mol^{-1}) and the natural phage λ repressor system (–2.0 kcal·mol^{-1}).[37,38]. Concomitantly, a surface plasmon resonance (SPR) assay was further validated quantitatively the cooperative effects of **Pip-HoGu** assembly (Table S3, Figure S7)[39].

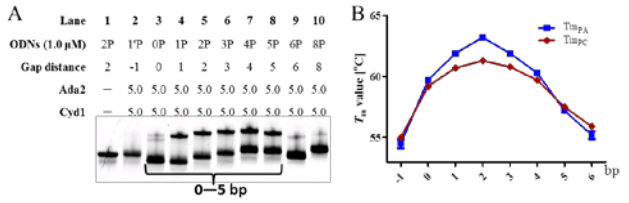


Figure 4. Mechanistic studies of cooperative binding. (A) Gel-shift behavior **Ada2–Cyd1**. (B) T_m profiles of all positive-binding sequences in the presence of **Ada1–Cyd1** (T_{mPA}, blue, same as Figure 2B) and **Ada2**–**Cyd1** (T_{mPB}, red).

Continually, we studied the influence of linker length on cooperative binding since the underlying mechanisms of gap-distance-dependent cooperativity (≤5 bp) are not totally clear[19]. For this, we used **Ada2**, which has a long, double mini-PEG linker (Figure 1). In the EMSA and T_m assay, **Ada2–Cyd1** showed similar cooperative patterns to **Ada1–Cyd1**, i.e., only those DNA sequences with gap distances of 0–5 bp could form cooperative complexes (Figure 4A, B). Specifically, **Ada2–Cyd1** showed lower stability for gap distances of 0–4 bp in the T_m assay. This demonstrated that an extra-long linker might destabilize the binding affinity of the complex over short gap distances. Interestingly, when the gap distance was extended to 5–6 bp, **Ada2–Cyd1** displayed slightly higher stability than **Ada1–Cyd1** because the longer and more flexible linker can reduce the tension of complex formation. In conclusion, the cooperative energy of **Pip-HoGu** was highly distance dependent and that gap distances of >5 bp diminish the cooperation, even when the linker region was long enough to allow the encounter of host–guest moieties. This can be explained by the hydrophobic and van der Waals interactions of Cyd–Ada[16,40].

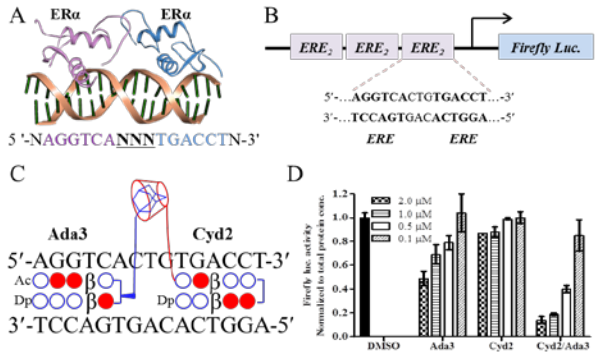


Figure 5. Cell-based assay of **Pip-HoGu**. (A) Crystal structure of ERα homodimer and DNA sequence. (B) Schematic diagram of ERE-driven luciferase in T47DK-BLuc cells. (C) The structural design of **Ada3** and **Cyd2** targeting ERE sites (Figure S8). (D) Luciferase activity assay after normalization to the total protein concentration.

Because we had compellingly demonstrated its cooperativity in several in vitro assay systems, we were encouraged to apply **Pip-HoGu** to a cell-based assay. The estrogen response element (ERE) is the specific target motif of the estrogen receptor α (ERα) homodimer, which induces significant downstream gene activation (Figure 5A)[41]. In ERα-positive, 17β-estradiol-stimulated T47D-KBluc cells that highly express luciferase after binding of three tandem ERα TF pairs, **Ada3**, **Cyd2**, and **Ada3**–**Cyd2** that bound to the ERE consensus half-site (5′-WGGWCW-3′) were tested for 48 h together with the delivery reagent endoporter (Figure 5B, C)[42]. The effects of PIPs were measured by luciferase activity normalized to total protein concentration[43,44]. Monotreatment with **Ada3** showed only moderate-to-weak inhibitory activity ($IC_{50} \geq 2$ μM) while very weak activity was observed for **Cyd2** (<20% inhibition) (Figure 5D). When cells were treated with a combination of **Ada3** and **Cyd2**, 4–5-fold enhancement of the inhibitory effect was observed, which strongly correlated with the cooperative interactions of the host–guest system. The use of endoporter ensured that **Ada3** and **Cyd2** reached homodimer-binding sites at the same time, and it is plausible that **Pip-HoGu** could also work well for heterodimer binding sites in the absence of endoporter[44]. This cell-based assay system demonstrated for the first time that the individual host–guest interactions of **Pip-HoGu** could effectively stabilize the PIP pair–DNA cooperative interaction and potently inhibit natural TF-pair binding in cells.

Cooperation between TF pairs is ubiquitous in cells[8]. Our prototype **Pip-HoGu** defines a general framework for mimicking cooperative TF pair–DNA interactions through the integration of programmable DNA binders and a host–guest system. In vitro assays showed that **Ada1–Cyd1** assemblies formed stable cooperative binding complexes with target DNA sequences with 0–5 bp gap distances. In essence, the cell-based assay demonstrated that our artificial **Ada3–Cyd2** assemblies formed highly stable cooperative complexes that competed with naturally occurring cooperative TF pair–DNA systems. Therefore, **Pip-HoGu** could be applied to the regulation of spatiotemporally variable gene expression patterns.

ASSOCIATED CONTENT

Supporting Information

Experimental details including synthesis, evaluation method and supporting data. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*hs@kuchem.kyoto-u.ac.jp

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant NO. JP16H06356, “Basic Science and Platform Technology Program for Innovative Biological Medicine by Japan Agency for Medical Research and Development (AMED)” , and “the Platform Project for Supporting Drug Discovery and Life Science Research funded by AMED”. We thank China Scholarship Council support Z. Y.

REFERENCES

(1) Takahashi, K.; Yamanaka, S. *Cell* **2006**, *126*, 663-676.
(2) Srivastava, D.; DeWitt, N. *Cell* **2016**, *166*, 1386-1396.
(3) Morgunova, E.; Taipale, J. *Curr. Opin. Struct. Biol.* **2017**, *47*, 1-8.
(4) Stampfel, G.; Kazmar, T.; Frank, O.; Wienerroither, S.; Reiter, F.; Stark, A. *Nature* **2015**, *528*, 147-151.
(5) Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1997**, *387*, 202-205.
(6) Dragulescu-Andrasi, A.; Rapireddy, S.; He, G.; Bhattacharya, B.; Hyldig-Nielsen, J. J.; Zon, G.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 16104-16112.
(7) Taniguchi, J.; Pandian, G. N.; Hidaka, T.; Hashiya, K.; Bando, T.; Kim, K. K.; Sugiyama, H. *Nucleic Acids Res.* **2017**, *45*, 9219-9228.
(8) Jolma, A.; Yin, Y.; Nitta, K. R.; Dave, K.; Popov, A.; Taipale, M.; Enge, M.; Kivioja, T.; Morgunova, E.; Taipale, J. *Nature* **2015**, *527*, 384-388.
(9) Aksoy, I.; Jauch, R.; Chen, J.; Dyla, M.; Divakar, U.; Bogu, G. K.; Teo, R.; Leng Ng, C. K.; Herath, W.; Lili, S.;

Hutchins, A. P.; Robson, P.; Kolatkar, P. R.; Stanton, L. W. *EMBO J.* **2013**, *32*, 938-953.

(10) Kamachi, Y.; Uchikawa, M.; Tanouchi, A.; Sekido, R.; Kondoh, H. *Genes Dev.* **2001**, *15*, 1272-1286.

(11) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559-561.

(12) Kurmis, A. A.; Yang, F.; Welch, T. R.; Nickols, N. G.; Dervan, P. B. *Cancer Res.* **2017**, *77*, 2207-2212.

(13) Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res.* **2004**, *32*, 2802-2818.

(14) Kawamoto, Y.; Sasaki, A.; Chandran, A.; Hashiya, K.; Ide, S.; Bando, T.; Maeshima, K.; Sugiyama, H. *J. Am. Chem. Soc.* **2016**, *138*, 14100-14107.

(15) Deplancke, B.; Alpern, D.; Gardeux, V. *Cell* **2016**, *166*, 538-554.

(16) Yu, G.; Jie, K.; Huang, F. *Chem. Rev.* **2015**, *115*, 7240-7303.

(17) Rodriguez, J.; Mosquera, J.; Garcia-Fandino, R.; Vazquez, M. E.; Mascarenas, J. L. *Chem. Sci.* **2016**, *7*, 3298-3303.

(18) Azuma, Y.; Imanishi, M.; Yoshimura, T.; Kawabata, T.; Futaki, S. *Angew. Chem. Int. Ed. Engl.* **2009**, *121*, 6985-6988.

(19) Ihara, T.; Uemura, A.; Futamura, A.; Shimizu, M.; Baba, N.; Nishizawa, S.; Teramae, N.; Jyo, A. *J. Am. Chem. Soc.* **2009**, *131*, 1386-1387.

(20) Machida, T.; Novoa, A.; Gillon, É.; Zheng, S.; Claudinon, J.; Eierhoff, T.; Imberty, A.; Römer, W.; Winssinger, N. *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 6762-6766.

(21) Zhou, X.; Su, X.; Pathak, P.; Vik, R.; Vinciguerra, B.; Isaacs, L.; Jayawickramarajah, J. *J. Am. Chem. Soc.* **2017**, *139*, 13916-13921.

(22) Morii, T.; Tanaka, T.; Sato, S.-i.; Hagihara, M.; Aizawa, Y.; Makino, K. *J. Am. Chem. Soc.* **2002**, *124*, 180-181.

(23) Lai, J.; Shah, B. P.; Garfunkel, E.; Lee, K.-B. *ACS Nano* **2013**, *7*, 2741-2750.

(24) Ueno, M.; Murakami, A.; Makino, K.; Morii, T. *J. Am. Chem. Soc.* **1993**, *115*, 12575-12576.

(25) Aizawa, Y.; Sugiura, Y.; Ueno, M.; Mori, Y.; Imoto, K.; Makino, K.; Morii, T. *Biochemistry* **1999**, *38*, 4008-4017.

(26) Blanco, J. B.; Dodero, V. I.; Vázquez, M. E.; Mosquera, M.; Castedo, L.; Mascareñas, J. L. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 8210-8214.

(27) Livengood, J. A.; Fechter, E. J.; Dervan, P. B.; Nyborg, J. K. *Front. Biosci.* **2004**, *9*, 3058-3067.

(28) Matsuoka, M.; Jeang, K.-T. *Nat. Rev. Cancer.* **2007**, *7*, 270-280.

(29) Guo, C.; Kawamoto, Y.; Asamitsu, S.; Sawatani, Y.; Hashiya, K.; Bando, T.; Sugiyama, H. *Bioorg. Med. Chem.* **2015**, *23*, 855-860.

(30) Ihara, T.; Takeda, Y.; Jyo, A. *J. Am. Chem. Soc.* **2001**, *123*, 1772-1773.

(31) Panjkovich, A.; Melo, F. *Bioinformatics* **2005**, *21*, 711-722.

(32) Yu, Z.; Taniguchi, J.; Wei, Y.; Pandian, G. N.; Hashiya, K.; Bando, T.; Sugiyama, H. *Eur. J. Med. Chem.* **2017**, *138*, 320-327.

(33) Kameshima, W.; Ishizuka, T.; Minoshima, M.; Yamamoto, M.; Sugiyama, H.; Xu, Y.; Komiyama, M. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 13681-13684.

(34) Hossain, M. A.; Hamasaki, K.; Takahashi, K.; Mihara, H.; Ueno, A. *J. Am. Chem. Soc.* **2001**, *123*, 7435-7436.

(35) Heddi, B.; Cheong, V. V.; Martadinata, H.; Phan, A. T. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 9608-9613.

(36) Sánchez, M. I.; Mosquera, J.; Vázquez, M. E.; Mascareñas, J. L. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 9917-9921.

(37) Distefano, M. D.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1179-1183.

(38) Ackers, G. K.; Johnson, A. D.; Shea, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 1129-1133.

(39) Asamitsu, S.; Li, Y.; Bando, T.; Sugiyama, H. *Chembiochem* **2016**, *17*, 1317-1322.

(40) Harada, A.; Takashima, Y.; Nakahata, M. *Acc. Chem. Res.* **2014**, *47*, 2128-2140.

(41) Boyer, M.; Poujol, N.; Margeat, E.; Royer, C. A. *Nucleic acids research* **2000**, *28*, 2494-2502.

(42) Wilson, V. S.; Bobseine, K.; Gray, J. L. E. *Toxicological Sciences* **2004**, *81*, 69-77.

(43) Nickols, N. G.; Szablowski, J. O.; Hargrove, A. E.; Li, B. C.; Raskatov, J. A.; Dervan, P. B. *Mol. Cancer Ther.* **2013**, *12*, 675-684.

(44) Summerton, J. E. *Annals of the New York Academy of Sciences* **2005**, *1058*, 62-75.

Supporting Information

Pip-HoGu, an artificial assembly with cooperative DNA recognition capable of mimicking transcription factor pair

Zutao Yu[†], Chuanxin Guo[†], Yulei Wei[†], Kaori Hashiya[†], Toshikazu Bando[†], Hiroshi Sugiyama[*†‡]

†Department of Chemistry, Graduate School of Science and ‡Institute for Integrated Cell-Material Sciences (iCeMS), Kyoto University, Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan

Table of Contents

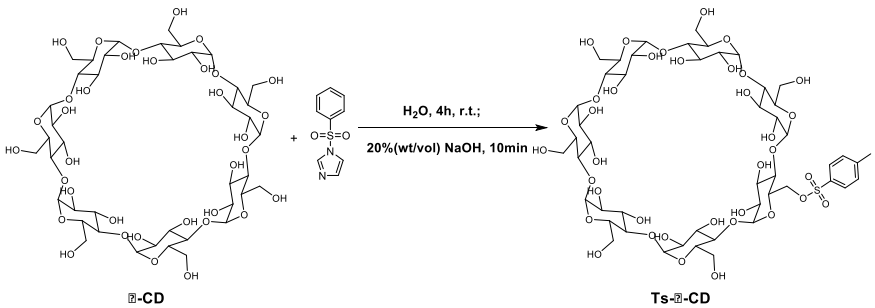
MATERIALS AND METHODS

General

The reagents for polyamide syntheses such as Fmoc-Py-OH, Fmoc-Im-OH, Fmoc-Py-Im-OH, and Im-CCl_3, solid supports (Fmoc-Py-oxime resin and Fmoc-β Ala-Wang resin), *O*-(1*H*-6-chlorobenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) and benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) were from HiPep Laboratories (Kyoto, Japan). Trifluoroacetic acid (TFA), 3,3′-diamino-*N*-methyldipropylamine, *N*,*N*-diisopropylethylamine (DIEA), dichloromethane (DCM), methanol, acetic acid (AcOH), 1-methyl-2-prrrolidone (NMP), and *N*,*N*-dimethylformamide (DMF) were obtained from Nacalai Tesque (Kyoto, Japan). Fmoc-D-Dab (Boc)-OH and Fmoc-NH-$dPEG_3$-COOH were obtained from Iris Biotech GmbH (Marktredwitz, Deutschland). Polyamide-chain assembly was performed on an automated synthesizer, PSSM-8 (Shimadzu, Kyoto, Japan). HPLC grade acetonitrile (Nacalai tesque) was used for both analytical and preparative HPLC. Water was prepared by a Milli-Q apparatus (Millipore, Tokyo, Japan). All chemicals were used as received. Analyses by reversed-phase RP-HPLC were carried out online LCMS (Agilent 1100 ion-trap mass spectrometer, HCT ultra, Bruker Daltonics, Yokohama, Japan), with analytical RP-HPLC columns, UV spectra were measured on a NanoDrop 2000c (Thermo Fisher Scientific).

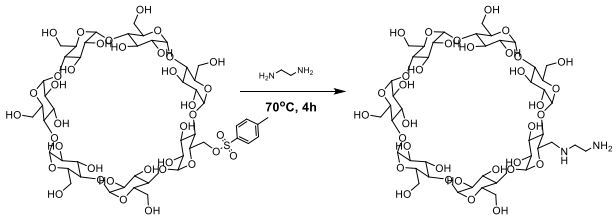
Synthesis of Ts-β-CD.

Weight out β-CD (7.0 g, 6.2 mmol) and 1-(*p*-toluenesulfonyl) imidazole (1.8 g, 8.0 mmol), and dissolved in 70 mL of water in a 200 ml conical flask containing a stir bar[1]. The suspension was stirred vigorously for 4 h at room temperature. Aqueous NaOH solution (20% wt/vol, 10 ml) was gradually added over 15 min. The mixture was stirred for an additional 10 min and filtered off the insoluble solid and collect the filtrate. Neutralize the filtrate to pH 7.0 with ammonium chloride (5.0 g, 0.1 mol) to induce precipitation. The precipitate was filtered off and washed with 3 × 10 mL of cold water and 2 × 10 mL of acetone. The product was dried in vacuum. The pure 6-*O*-monotosyl-β-CD (Ts-β-CD) was dried overnight in oil vacuum and obtained as white powder (2.8g, 35% yield). ESI-TOF-MS: *m*/*z* calcd for $C_{49}H_{77}O_{37}S^+$ $[M+H]^+$: 1289.3824; found: 1289.2265, 1290.2354.



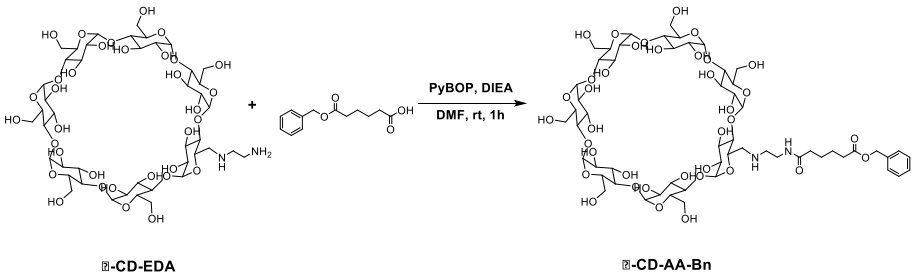
Synthesis of β-CD-EDA

2.8 g of Ts-β-CD was reacted with excess amount of EDA (30 ml) at 75℃ for 4 h. After the reaction was completed, the mixture was allowed to cool to room temperature, and then put the solution in ice water[2]. Gradually added 30 ml of cold acetone and cooled down the solution to 0℃. The precipitate was repeatedly dissolved in 15 ml of water-methanol mixture, and poured into 30 ml of acetone several times for the removal of unreacted EDA. The sample obtained was dried at vacuum for 24 h and pure β-CD-EDA was obtained as white powder (1.5 g, 59% yield). ESI-TOF-MS: *m*/*z* calcd for $C_{44}H_{77}N_2O_{34}^+$ $[M+H]^+$: 1177.4358; found: 1177.2965, 1178.3079.



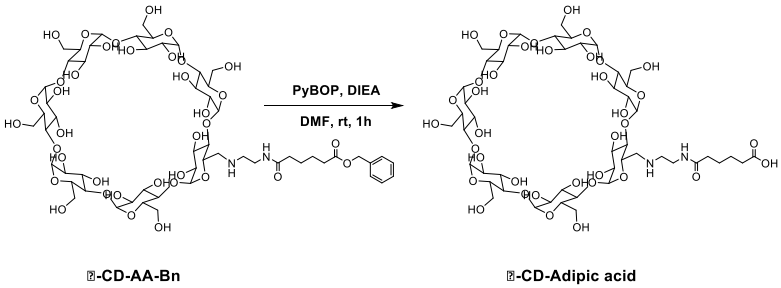
Synthesis of β-CD-AA-Bn

To a 25 ml conical flask, there was sequentially added 52 mg 6-(benzyloxy)-6-oxohexanoic acid (AA-Bn), 20 mg PyBOP and 15 ml DMF. Then 13 ul DIEA was injected into the mixture and stirred for 15 min for carboxyl acid group activation at room temperature. Added 250 mg β-CD-EDA and stirred vigorously at room temperature for 1h. The precipitate was repeatedly dissolved in 15 ml of water-methanol mixture and poured into 30 ml of acetone several times. The sample obtained was dried at vacuum for 24 h and pure EDA-β-CD was obtained as white powder (187 mg, 63% yield). ESI-TOF-MS: *m*/*z* calcd for $C_{57}H_{91}N_2O_{37}^+$ $[M+H]^+$: 1395.5301; found: 1395.5491, 1396.5601.



Synthesis of β-CD-Adipic acid

187 mg of β-CD-AA-Bn and 1.8 mg Pd/C were added to 15 ml methanol in 25 ml flask, then put it under hydrogen reduction system for 2 h. Removed Pd/C by double filtration and then pure into ether followed by AcOEt. Then white solid was obtained by centrifugation (174 mg, 99% yield). ESI-TOF-MS: *m*/*z* calcd for $C_{50}H_{85}N_2O_{37}^+$ $[M+H]^+$: 1305.4831; found: 1305.3771, 1306.3610.



-CD-AA-Bn　　-CD-Adipic acid

Polyamide Fmoc coupling procedure

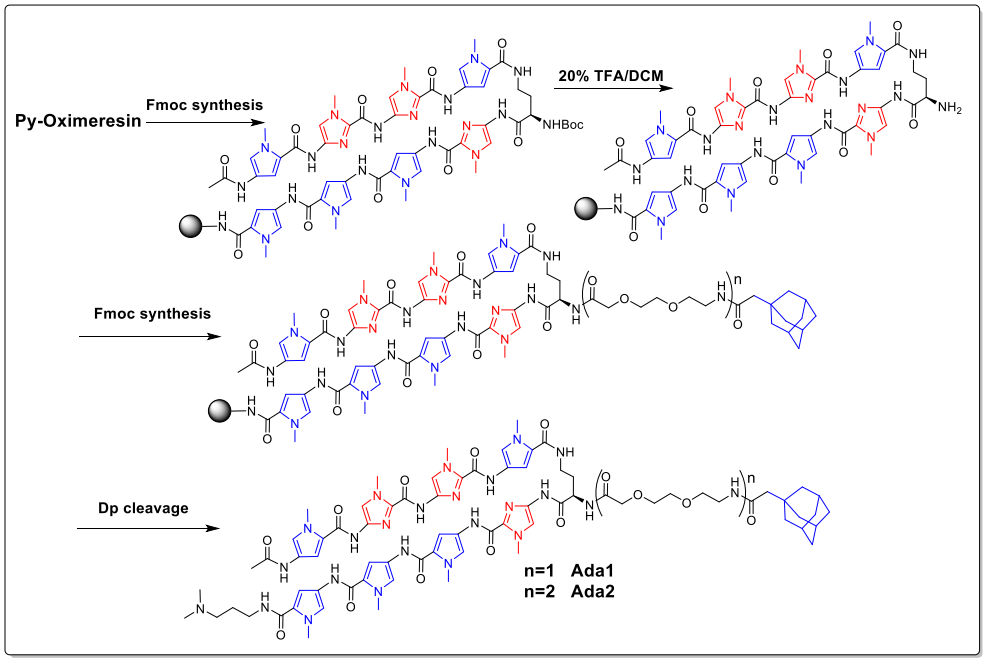
Polyamides were prepared using PSSM-8 peptide synthesizer (Shimadzu, Kyoto) with a computer-assisted operation system at 43 mg of Fmoc-Pyrrol-oxime resin and β Ala-Wang resin (ca. 0.42 mmol/g, 100~200 mesh) by Fmoc solid-phase chemistry[3,4]. Reaction cycles were as follows: deblocking step for 4 min x 2, 20% piperidine in DMF; coupling step for 60 min, corresponding carboxylic acids, HCTU (88 mg), diisopropylethylamine (DIEA) (36 μL), 1-methyl-2-pyrrolidone (NMP); washing step for 1 min x 5, DMF. Each coupling reagents in steps were prepared in NMP solution of Fmoc-Py-COOH (77 mg), Fmoc-Im-COOH (77 mg), Fmoc-Py-Im-COOH (70 mg), Fmoc-β-COOH (66 mg), Fmoc-γ-COOH (69 mg) and Fmoc-mini PEG-COOH (69 mg). All other couplings were carried out with single-couple cycles with stirred by N_2 gas bubbling. Typically, resin (40 mg) was swollen in 1 mL of NMP in a 2.5-mL plastic reaction vessel for 30 min. 2-mL plastic centrifuge tubes with loading Fmoc-monomers with HCTU in NMP 1 mL were placed in programmed position. All lines were washed with NMP after solution transfers. After the completion of the synthesis by the peptide synthesizer, the resin was washed with DMF (1 mL × 2), methanol (1 mL × 2), and dried in a desiccator at room temperature in vacuo.

Resin cleavage and purification procedure

The resulting polyamide-oxime resin was cleaved from the solid support with *N*,*N*-dimethyl-1,3-propyldiamine for 3 h at 45 °C. Polyamide-β Ala-Wang resin was cleaved from the solid support with 95% TFA, 2.5% triisopropylsilane, and 2.5% water for 30 min at room termperature. Resin was filtered off, and the resulting liquor was treated with diethyl ether. The precipitated crude polyamide was washed three times with diethyl ether and analyzed by RP-HPLC. Crude polyamides were purified on a preparative column at 40°C. The purified peptides were assessed by the LC-MS system.

Synthesis of Ada1

Polyamide synthetic procedure has been described above. The condition of Boc- deprotection is 20% TFA in DCM, 30 min at room temperature. The resin was washed by DMF, DCM, and dried in vacuum. Then Fmoc chemistry was conducted for the coupling of mini-PEG linker. After deprotection of Fmoc group, 1-Adamantaneacetic acid (40 mg) was coupled by Fmoc method. The resin cleavage and compound purification procedure have been described above. **Ada1** was obtained as a white powder. MALDI-TOF MS: *m*/*z* calcd for $C_{74}H_{97}N_{24}O_{14}^{+}$ $[M+H]^{+}$: 1545.7616; found: 1545.107.. HPLC: t_R=12.667 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–20 min). (Mass data was attached in the bottom)

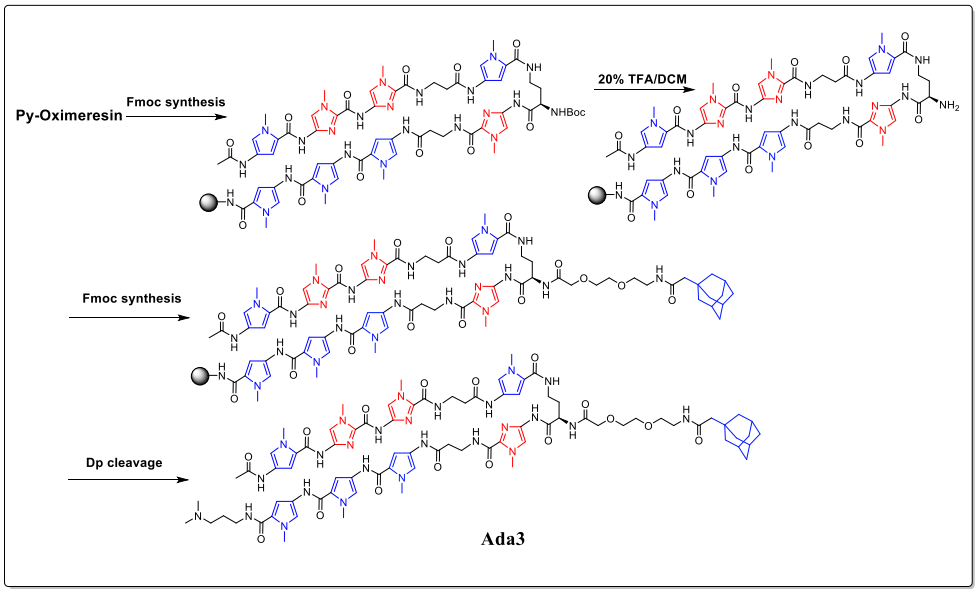


Synthesis of Ada2

The synthetic procedure was same as **Ada1**, except double coupling step of mini-PEG linker. **Ada2** was obtained as a white powder. ESI-TOF MS: *m/z* calcd for $C_{80}H_{110}N_{25}O_{17}^{3+}$ $[M+3H]^{3+}$: 564.2837; found: 564.2218, 564.5593, 564.9067; $C_{80}H_{109}N_{25}O_{17}^{2+}$ $[M+2H]^{2+}$: 845.9216; found: 845.8101, 846.3149, 846.8141. HPLC: t_R=12.775 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–20 min). (Mass data was attached in the bottom)

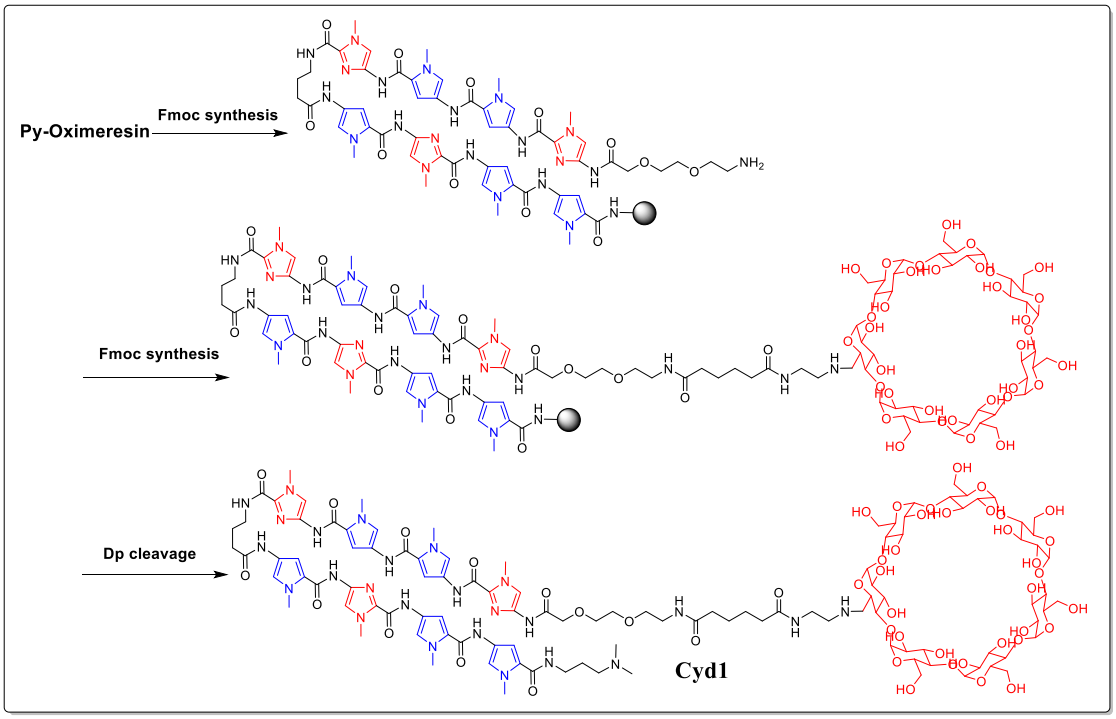
Synthesis of Ada3

Ada3 was obtained as a white powder. MALDI-TOF MS: *m/z* calcd for $C_{80}H_{105}N_{26}NaO_{16}$ $[M+Na]^{+}$: 1709.876; found: 1709.975. HPLC: t_R=20.325 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–40 min).



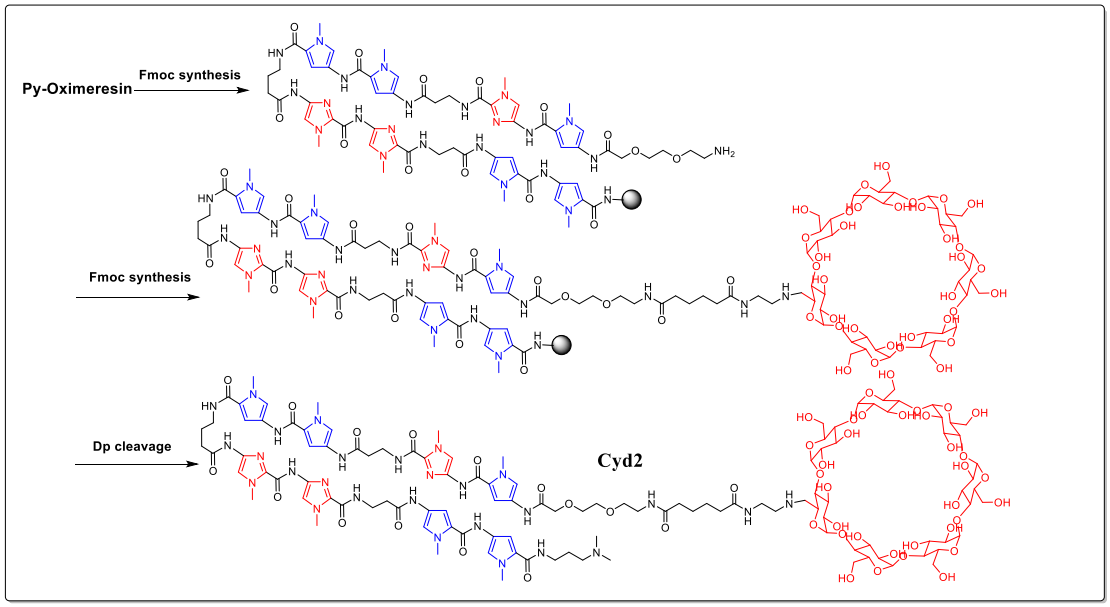
Synthesis of CyD1

Polyamide synthetic procedures have been described above. The coupling condition of linker is adipic acid (12.5 mg), HCTU (74 mg), DIEA (22 μL), NMP (1000 μL), 60 min. Then washed by DMF, DCM, and dried in vacuum. The resin was activated by PyBOP for 30min, then EDA-β-CD conjugated to the polyamides and stirred for 2 h, and washed by above mentioned solvent. The resulting polyamide-resin was cleaved from the solid support with *N*,*N*-dimethyl-1,3-propyldiamine for 3 h at 45 °C. Resin was filtered off, and the resulting liquor was treated with DCM. The precipitated crude polyamide was washed three times with DCM and analyzed by RP-HPLC. Crude polyamides were purified on a preparative column, HiPep-Intrada, at 40 °C. The purified peptides were assessed by the LC-MS system. **CyD1** was obtained as a white powder. ESI-TOF MS: *m*/*z* calcd for $C_{110}H_{163}N_{25}O_{48}^{4+}$ $[M+4H]^{4+}$: 650.7779; found: 650.4527, 650.7227; $C_{110}H_{162}N_{25}O_{48}^{3+}$ $[M+3H]^{3+}$: 867.3679; found: 866.9463, 867.2734, 867.6157. HPLC: t_R=9.500 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–20 min). (Mass data was attached in the bottom)



Synthesis of Cyd2

Cyd2 was obtained as a white powder. MALDI-TOF MS: *m*/*z* calcd for $C_{116}H_{168}N_{27}NaO_{50}^{+}$ $[M+Na]^{+}$: 2763.1404; found: 2764.437. HPLC: t_R=14.775 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–40 min).



Synthesis of PIP1

PIP1 was obtained as a white powder. ESI-TOF MS: *m/z* calcd for $C_{61}H_{78}N_{23}O_{12}^{+}$ $[M+H]^{+}$: 1324.6195; found: 1324.3297, 1325.3289; $C_{61}H_{78}N_{23}O_{12}^{2+}$ $[M+2H]^{2+}$ found: 662.6921, 663.1908. HPLC: t_R=11.675 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–20 min). (Mass data was attached in the bottom)

Synthesis of PIP2

PIP2 was obtained as a white powder. MALDI-TOF MS: *m/z* calcd for $C_{62}H_{80}N_{23}O_{13}^{+}$ $[M+H]^{+}$: 1354.6300; found: 1353.893; $C_{61}H_{78}N_{23}O_{12}^{+}$ $[M+Na]^{+}$ found: 1375.911. HPLC: t_R=11.208 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–20 min). (Mass data was attached in the bottom)

Thermal Denaturation Analyses

Thermal denaturation analyses of the polyamide–DNA complex were performed on a V-650 spectrophotometer (JASCO) having a cell path length of 1 cm equipped with a thermocontrolled PAC-743R cell changer (JASCO) and a refrigerated and heated circulator F25-ED (Julabo) as described[3]. The sequences of the DNAs used were purchased from Sigma-Aldrich. The analysis buffer is as follows: the aqueous solution of 10 mM sodium chloride and 10 mM sodium cacodylate at pH 7.0 containing 0.25% v/v DMSO. The final concentrations of polyamides and dsDNA were 7.5 μM and 2.5 μM, respectively (3:1 stoichiometry). Denaturation profiles were recorded at λ = 260 nm from 25 to 95 °C at a rate of 1.0 °C/min, and melting temperatures were measured as the maximum of the first derivative of the profiles. Reported values were the averages of at least three measurements.

Electrophoretic mobility shift assay (EMSA)

Preparation loading mixture[5]. The sequences of the DNAs used were purchased from Sigma-Aldrich. The analysis buffer is as follows: the aqueous solution of 10 mM sodium chloride and 10 mM sodium cacodylate at pH 7.0 containing 0.25% v/v DMSO. The final concentrations of polyamides and dsDNA were clarified in the manuscript. Gel Loading Dye, Purple (6X no SDS B7025S, New England Biolab).

Preparation of gels. In a clean glass beaker the following reagents were mixture in the given order (10 ml system, reagent volume doubled for 20 ml system). 5.25 mL MiliQ, 1 mL 10× TBE, and 3.75 mL of 40% Acrylamide/Bis Solution (29 : 1), followed by gas-removing to ensure the removal of all air bubbles. Then 90 μL APS (10% w/w in MiliQ) and 100 μL TEMED (10% v/v in MiliQ) were then added to the mixture and mixed properly before pouring it gently along parallel glass plates. Sufficient time was given for polymerization (20 min).

Electrophoresis. A pre-run of the gels was performed prior to loading. Care was taken to see that the gel were properly immersed in 1× Tris-Borate-EDTA buffer (TBE buffer) and the loading wells were free from any air bubbles. The wells were washed after the pre-run. Instrument settings: 120 V for 30 min at 4°C. 4 μL of the loading mixture was then loaded onto the wells. Pre-run again at 120 V for 30 minutes at 4°C. Then gel running as the instrument settings: 180 V for 160 min at 4°C.

Analysis of gels. The bands were stained with SYBR gold (10000× concentration in DMSO, from Thermofisher) and quantified with a FujiFilm FLA-3000G fluorescent imaging analyzer.

SPR Binding Assays

The SPR assays were performed using a BIACORE X instrument as previously described[6]. Biotinylated hairpin DNA purchased from JBioS, whose sequence was 5′-biotin-AACTTAGGCTAATGACGTATATGTTTTCATATACGTCATTAGCCTAAGTT-3′, was immobilized on streptavidin-coated sensor chip SA to obtain the desired immobilization level (approximately 1100 RU rise). The assays were carried out using HBS-EP (10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% v/v Surfactant P20), purchased from GE Healthcare,

with 0.1% DMSO at 25 °C. A series of sample solutions with various concentrations were prepared in the buffer with 0.1% DMSO and injected at a flow rate of 20 mL/min. To calculate the association rate (k_a), dissociation rate (k_d), and dissociation constant (K_D), data processing was performed with 1:1 binding with mass a transfer model using BIA evaluation 4.1 program.

Cell culture and Cell treatment

T47D-KBluc cell line used was purchased directly from American Type Culture Collection (ATCC) and no subsequent authentications were done by the authors[7]. Cells were grown in RPMI-1640 (Gibco) held at 37°C in 5% CO_2. Media was supplemented with 10% FBS (Sigma, unheated) and 1% penicillin/streptamycin. Before one week of compound treatment, media was supplemented with 10% FBS (Sigma, unheated) without penicillin/streptomycin.

T47D-KBluc cells were plated at 2×10^5 cells per well in 24-well plates, incubated in phenol red-free RPMI 1640 medium growth media (Gibco) containing 10 nmol/L E2 supplemented with 10% hyclone™ FBS (GE healthcare, unheated)[8]. After 24 hours, plates were replenished with same medium dosed with 0.1-2 μmol/L polyamides and then add 3 μl Endoporter (Gene Tools) to the concentration of 6 μmol/L. The cells were then incubated for 48 hours[9].

BCA protein assay and luciferase assay

Cell lyses were obtained with 80 μl/well luciferase cell culture lysis reagent (Promega). Then cell lyses were analyzed using both Pierce™ BCA protein assay (Thermo) and luciferase assay system (Promega) according to the manufacturers' instructions. Luciferase assay data were normalized with BCA protein assay[10].

Statistical analysis

Results for continuous variables were presented as the mean ± standard error. Two-group differences in continuous variables were assessed by the unpaired t-test. Statistical analysis was performed by comparing treated samples with untreated controls. The statistical analyses were performed using GraphPad Prism 5.

Computational experiments

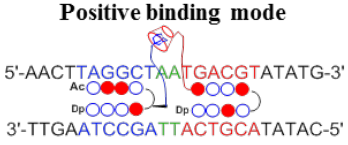
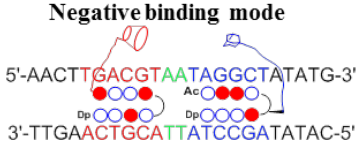
Software BIOVIA Discovery Studio version 17.1 from Accelrys was used for this purpose[11]. The primary geometric PIP structure, β-CD, and adamantine were taken from the crystal structure 3OMJ[12], 4D5B[13], and 1LWL[14] respectively. The starting structures of DNA, **Ada1**, and **Cyd1** were modified and energy minimized using the tools in the software. After energy minimization without any constrains, the distance of hydrogen bond between PIPs and DNA are ranging from 1.928 to 3.328 Å. The following table list one example of Distance Monitoring data (2 bp gap distance, Figure S1):

Name	Visible	Color	Parent	X	Y	Z	Distance
Distance40	Yes	0 255 0	<TreeRoot>	-3.76078	7.09167	-5.21038	2.314
Distance50	Yes	0 255 0	<TreeRoot>	-3.18661	9.64179	-8.72648	2.642

Distance60	Yes	0	255	0	<TreeRoot>	-0.939881	11.5555	-12.3476	2.047
Distance70	Yes	0	255	0	<TreeRoot>	-0.298467	11.8243	-14.8366	2.821
Distance80	Yes	0	255	0	<TreeRoot>	1.87322	10.884	-15.7153	2.814
Distance90	Yes	0	255	0	<TreeRoot>	3.47262	9.02065	-19.3201	1.928
Distance11	Yes	0	255	0	<TreeRoot>	1.16246	10.2712	-22.2429	2.014
Distance12	Yes	0	255	0	<TreeRoot>	0.801845	11.1992	-20.2019	2.578
Distance13	Yes	0	255	0	<TreeRoot>	-1.45549	10.4291	-19.2067	3.078
Distance14	Yes	0	255	0	<TreeRoot>	-3.37622	9.14564	-15.4857	2.652
Distance15	Yes	0	255	0	<TreeRoot>	-4.34768	6.72967	-11.6213	3.058
Distance16	Yes	0	255	0	<TreeRoot>	-4.9555	6.03272	-9.09647	3.328
Distance17	Yes	0	255	0	<TreeRoot>	-3.36976	4.24082	-8.41492	3.304
Distance18	Yes	0	255	0	<TreeRoot>	-1.66559	10.2419	22.9747	3.079
Distance19	Yes	0	255	0	<TreeRoot>	0.625469	11.132	18.967	3.301
Distance20	Yes	0	255	0	<TreeRoot>	3.37375	10.2939	15.622	2.531
Distance21	Yes	0	255	0	<TreeRoot>	4.31715	7.83195	12.1853	2.673
Distance22	Yes	0	255	0	<TreeRoot>	3.79255	5.27068	8.63684	2.704
Distance23	Yes	0	255	0	<TreeRoot>	5.12323	7.09738	9.85063	2.645
Distance24	Yes	0	255	0	<TreeRoot>	3.02919	10.7066	9.23694	2.29
Distance25	Yes	0	255	0	<TreeRoot>	2.70262	11.6067	11.6097	2.641
Distance26	Yes	0	255	0	<TreeRoot>	0.251417	11.73	12.1004	2.751
Distance27	Yes	0	255	0	<TreeRoot>	-0.151984	12.2931	15.0042	2.869
Distance28	Yes	0	255	0	<TreeRoot>	-1.99701	10.8087	15.4324	2.667
Distance29	Yes	0	255	0	<TreeRoot>	-3.61237	8.14586	18.8926	2.075

SUPPORTING TABLES AND FIGURES

Table S1. Results of T_m assay

Gap distance (n)	Positive binding mode 			Negative binding mode 			$\Delta\Delta T_m$ [℃]
	ODNs	T_m [℃]	T_{mPA} [℃]	ODNs	T_m [℃]	T_{mNA} [℃]	
-1 bp	1'P	45.7	54.5	1'N	45.2	54.8	-0.8
0 bp	0P	46.2	59.7	0N	45.7	57.0	2.1
1 bp	1P	45.5	61.9	1N	44.8	56.1	5.1
2 bp	2P	45.8	63.2	2N	45.4	55.7	7.2
3 bp	3P	46.6	61.9	3N	46.1	56.9	4.6
4 bp	4P	46.1	60.3	4N	45.7	55.7	4.2
5 bp	5P	45.0	57.2	5N	44.5	55.0	1.7
6 bp	6P	44.8	55.2	6N	44.2	54.9	-0.3

*ΔT_m = T_m (ODNs/PIPs) - T_m (ODNs); $\Delta\Delta T_m$ = ΔT_{mP} - ΔT_{mN}. Error bars are ranging from 0.25-0.7 °C indicating standard deviations of three replicates.

Table S2. Results of Tm assay of mismatch sequence

ODNs	T_m [℃]	T_{mPA} [℃]	ΔT_{mA} [℃][a]	$\Delta\Delta T_m$ [℃]
ODN2P	45.8	63.2	17.4	8.4
ODN2PM	43.9	52.9	9.0	

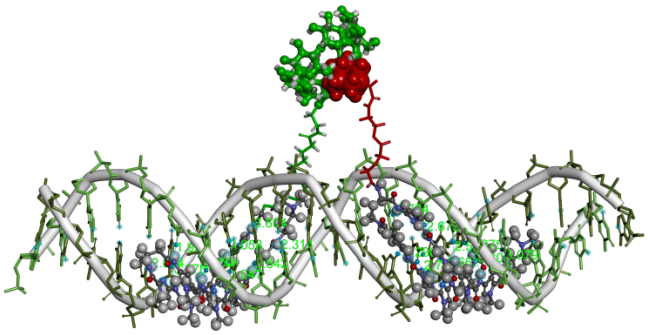
ODN2P: 5'-AACTTAGGCTAATGACGTATAT-3'

ODN2PM: 5'-AACTTAGGCTAATGATGTATAT-3'

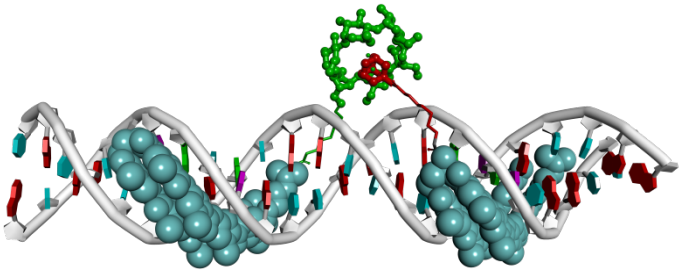
[a] ΔT_m = T_m (ODNs/PIPs) - T_m (ODNs); $\Delta\Delta T_m$ = ΔT_{mP} - ΔT_{mPM}. Error bars are ranging from 0.2-0.65 °C indicating standard deviations of three replicates.

Figure S1. Computational modeling structure of positive DNA sequences with Ada1/Cyd1 assemblies.

2 bp gap distance
(Retaining structural constrain and distance, dihedral angle monitor)



5 bp gap distance
(Simplified model structure)



6 bp gap distance
(Simplified model structure)

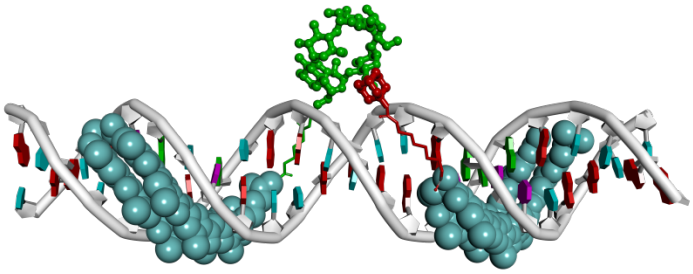
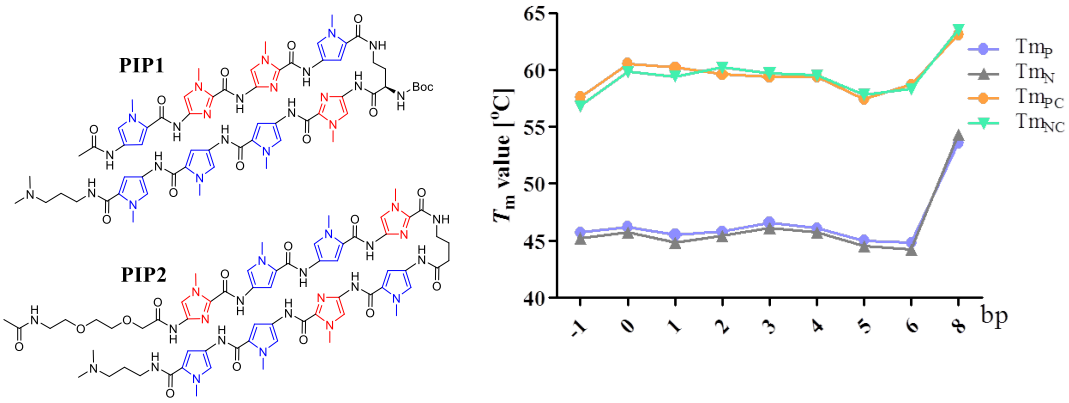


Figure S2. Results of T_m assay of PIP1–PIP2



(A) Chemical structures of PIP1 and PIP2; (B) T_m profiles of positive ODNs (T_{mP}, light blue line), negative ODNs (T_{mN}, gray line), positive ODNs/PIP1–PIP2 (T_{mPC}, brown line), and negative ODNs/ PIP1–PIP2 (T_{mNC}, green line). *Error bars are ranging from 0.25-0.7 °C indicating standard deviations of three replicates.

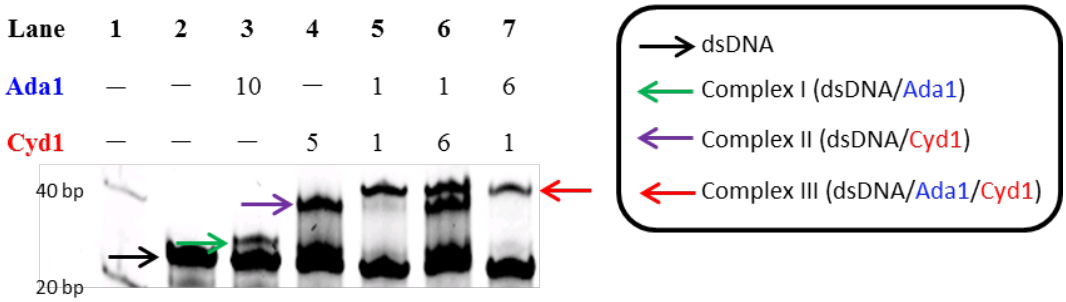
ODN8P: 5'-AATTAGGCTGGAATTCCTGACGTATA-3';

ODN8N: 5'-AATTGACGTGGAATTCCTAGGCTATA-3'.

Here only show forward DNA strands with complimentary DNA strands omitted.

At the gap distance of **–**1 bp in positive- and negative-binding sequences, the T_m value was far lower than others, which were in consistent with spatial incompatibility upon two PIPs binding. The huge increase of T_m value at the gap distance of 8 bp can be explained by the insertion of multiple G/C in the gap region. One thing should be noticed here that, the conjugation of host-guest moieties in PIPs impair its binding affinity to the target DNA sequences to a reasonable extent, which could be observed by comparing T_{mNA} (Negative-ODNs/**Ada1–Cyd1**) and T_{mNC} (Negative-ODNs/**PP1–PP2**). Therefore, in the control experiment with mixture of PIP1 and PIP2, there was no significant difference of thermal stability between positive- and negative-binding sequences (Figure S2)[15].

Figure S3. Up-shifted band of ODN2P with Ada1, Cyd1, and Ada1/Cyd1 respectively.



*Discrimination of up-shifted band of ODN2P with **Ada1** (green arrow), **Cyd1** (purple arrow), and **Ada1/Cyd1** (red arrow) in EMSA. Black arrow indicates ODN2P duplex. All dsDNA concentration: 1 μM.

Figure S4. Up-shifted band of ODN2P of positive binding mode and ODN2N of negative binding mode with Ada1/Cyd1 respectively.

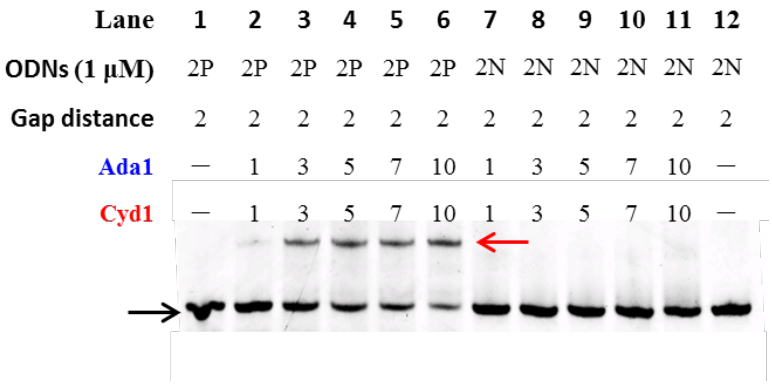
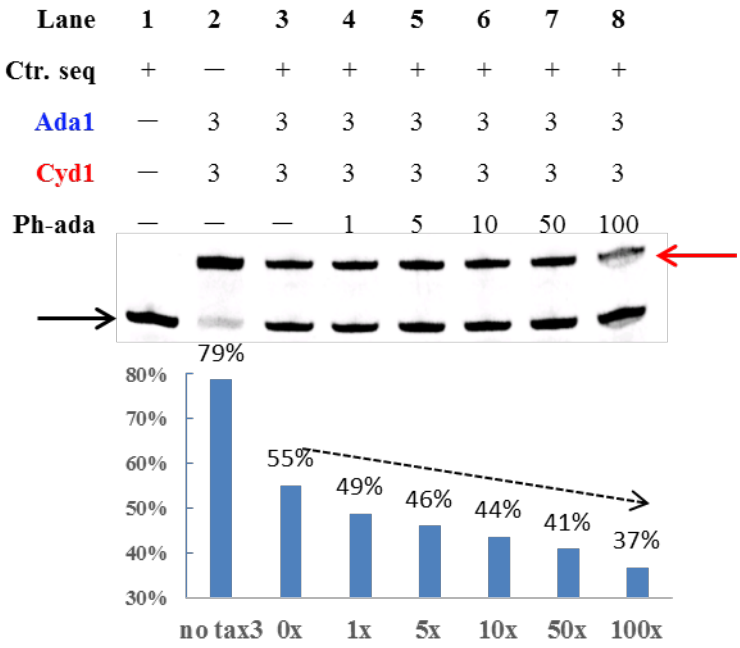
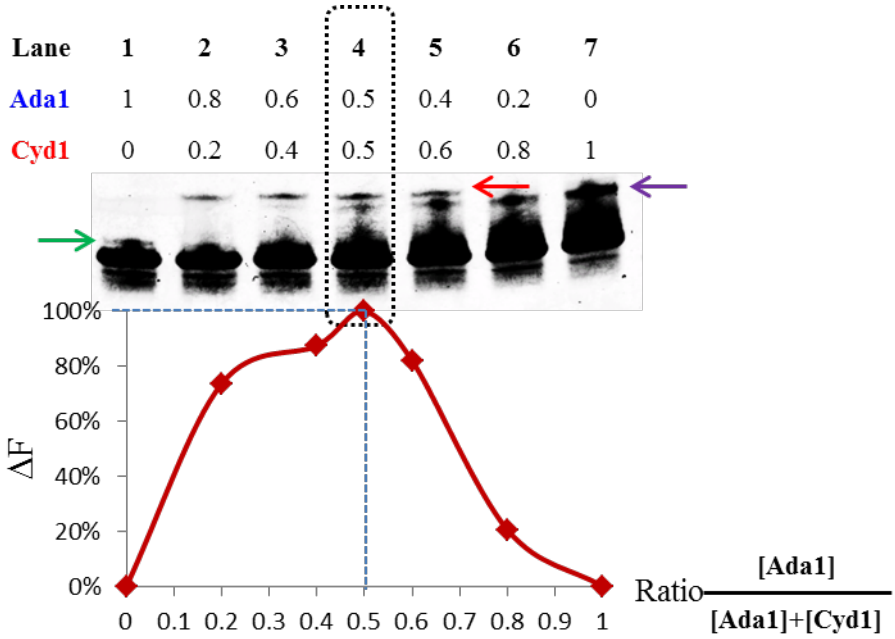


Figure S5. Competitive assay of Ada1/Cyd1/ODN2P with Phenyladamantane in EMSA.



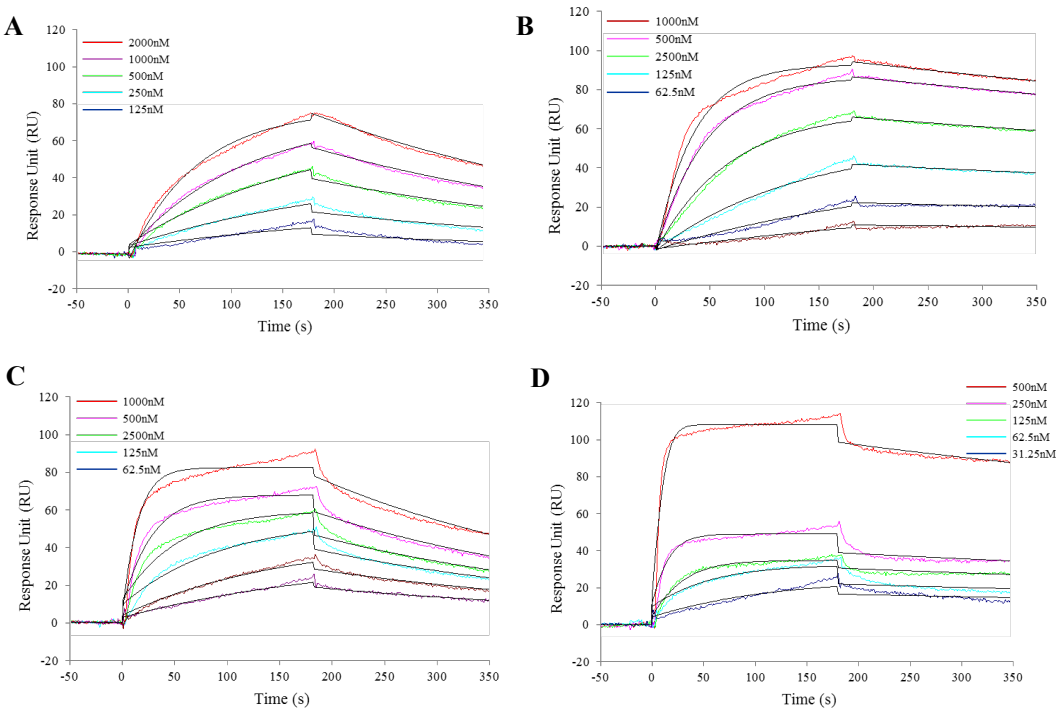
*ODN2P concentration: 1 μM. **Ada1**/**Cyd1** concentration: 3 μM. Ph-ada concentration is from 1 μM, 5 μM, 10 μM, 50 μM, and 100 μM. Ctr.seq is one competitive sequence for **Ada1** binding site, 10 μM.

Figure S6. Job plot of Ada1/Cyd1/ODN2P in EMSA.



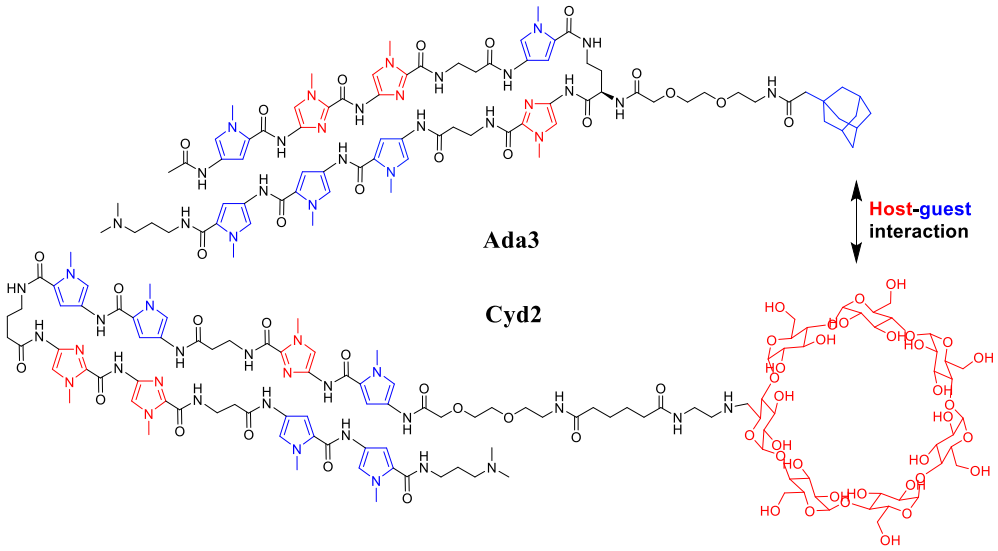
*The total concentration of conjugates was constant at 1×10^{-6} M by preparing a total of 12 μL mixture of solutions.. ΔF is the percentage of band shift in fluorescence intensity normalized to lane 4. X is the mole fraction of **Ada1** to **Cyd1**. The highest band shift point at 0.5 fraction of [**Ada1**]/[**Ada1**]+[**Cyd1**] indicates that the stoichiometry of binding is of 1:1 ratio of two conjugates per mole of DNA strands. ODN2P concentration: 1 μM.

Figure S7. SPR binding assays conducted to evaluate the binding properties of Ada1 and Cyd1.



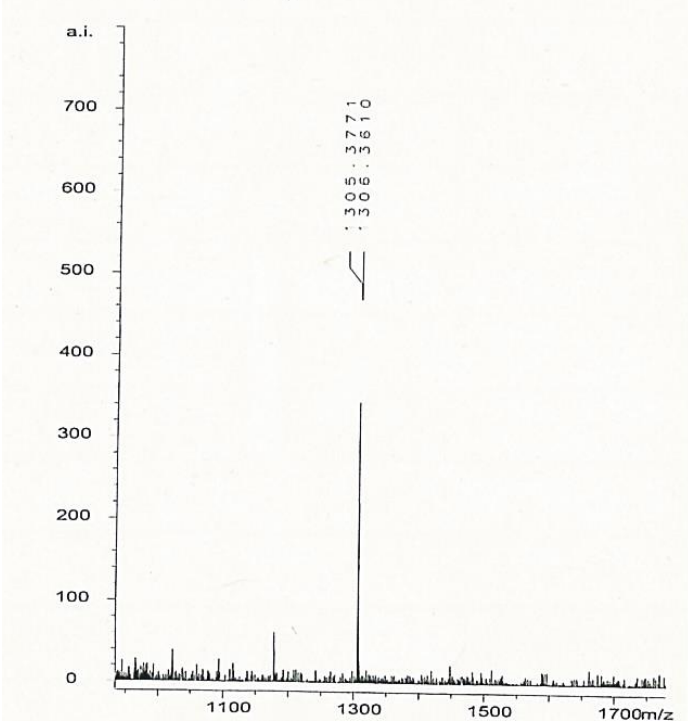
*SPR sensorgrams for interactions of **Ada1** (A), **Ada1** + fixed concentration of **Cyd1** (B), **Cyd1** (C), and **Cyd1** +fixed concentration of **Ada1**. The concentrations were showed in figure. Extensive concentration of **Ada1** and **Cyd1** were dissolved in HBS-EP buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, and 0.005 % surfactant P20) with 0.1% DMSO. These solutions were passed over a 5'-biotinylated DNA hairpin, which contain the binding sites of **Ada1** and **Cyd1,** immobilized on a sensor chip through a biotin-avidin system. Under these conditions, the binding affinities were evaluated and by applying 1:1 binding model with mass transfer to each sensorgram, the association rate constant k_a, dissociation rate constant k_d, and disassociation equilibrium constant K_D could be determined. Kinetic constants were calculated from the surface plasmon resonance sensorgrams for the interaction of polyamide with the DNA immobilized on the surface of a sensor chip SA (Biacore assays). K_D, dissociation equilibrium constant; k_a, association rate constant; k_d, dissociation rate constant.

Figure S8. Chemical structures of Ada3 and Cyd2.

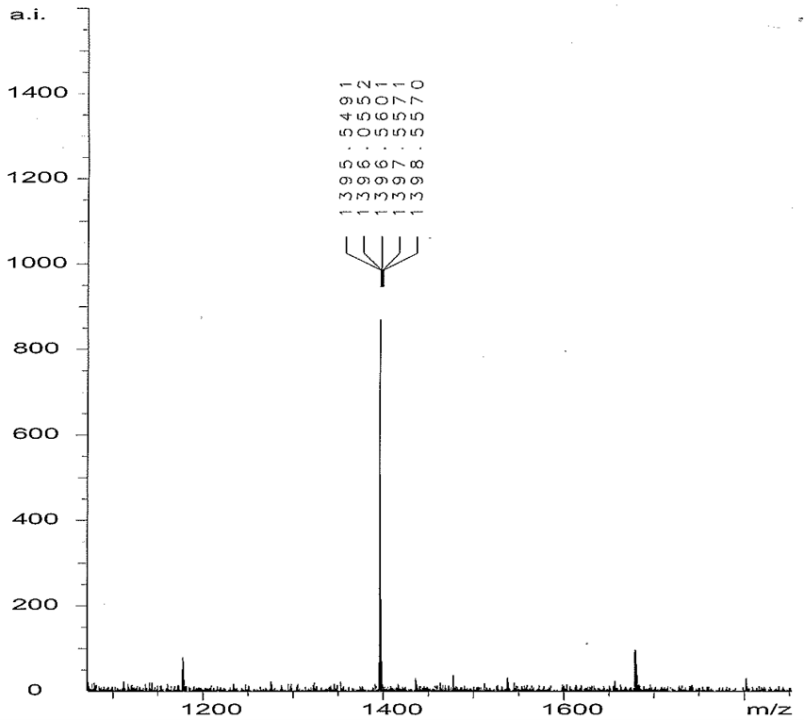


MS data

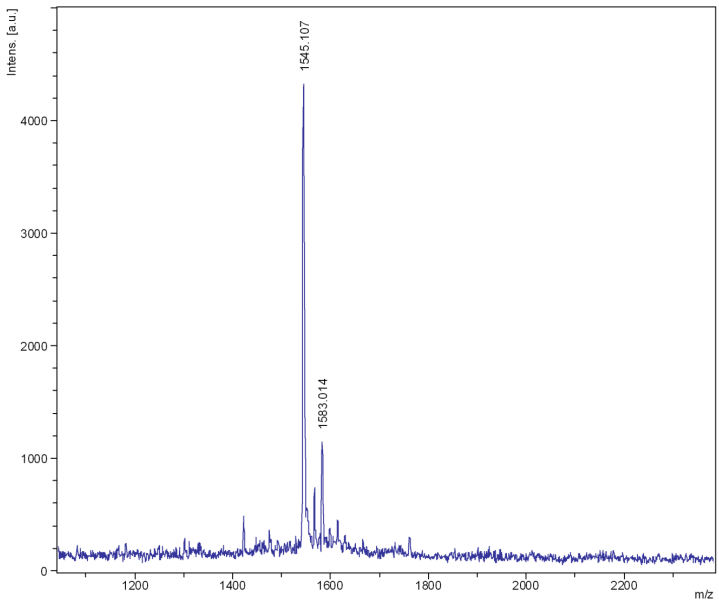
β-CD-AA-Bn. was obtained as white powder. ESI-TOF-MS: *m*/*z* calcd for $C_{57}H_{91}N_2O_{37}^+$ $[M+H]^+$: 1395.5301; found: 1395.5491, 1396.5601.



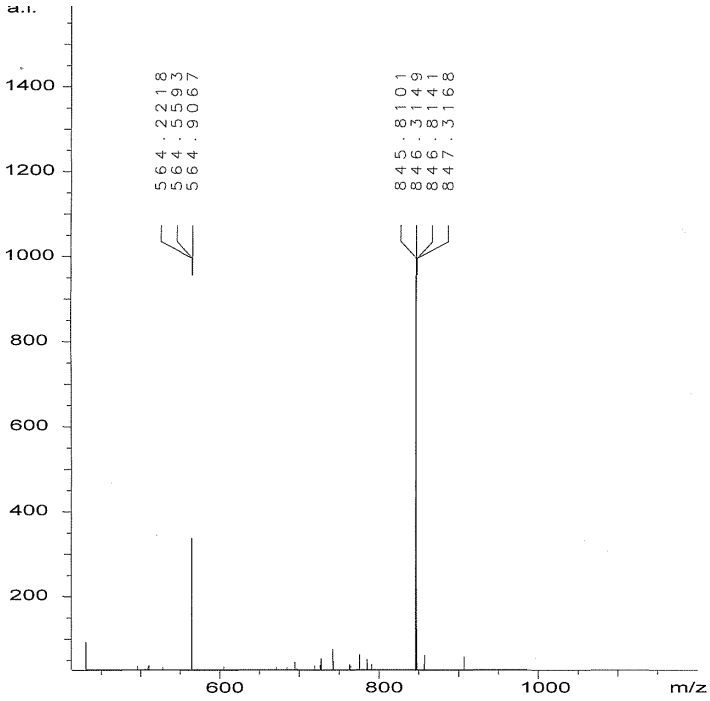
β-CD-Adipic acid was obtained as white powder. ESI-TOF-MS: *m*/*z* calcd for $C_{50}H_{85}N_2O_{37}^+$ $[M+H]^+$: 1305.4831; found: 1305.3771, 1306.3610.



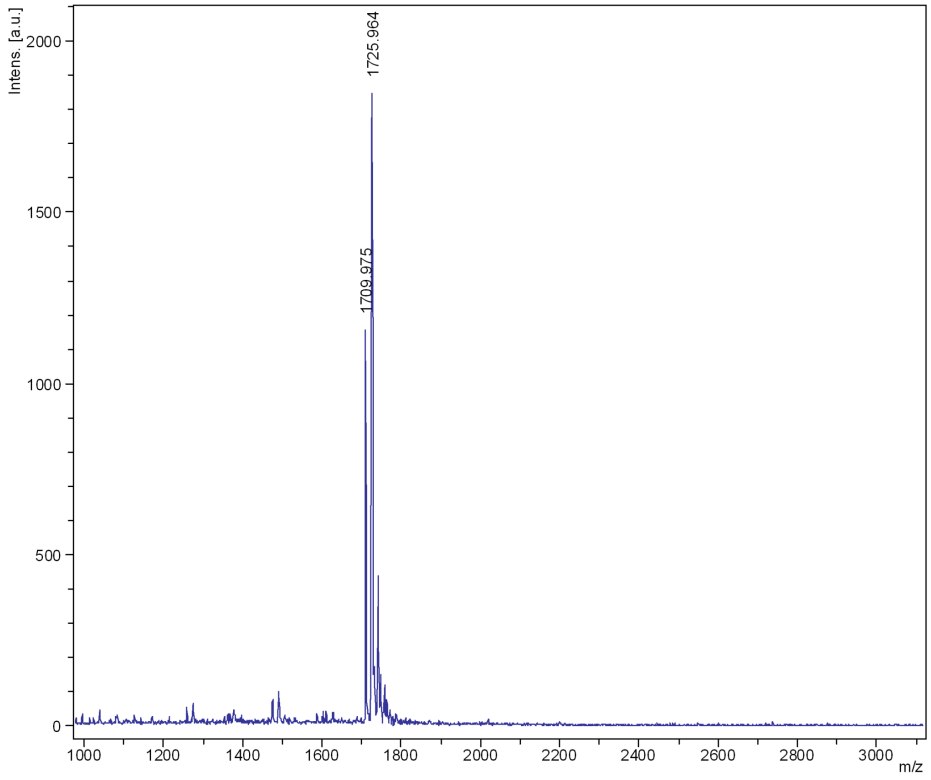
Ada1 was obtained as a white powder. MALDI-TOF MS: *m/z* calcd for $C_{74}H_{97}N_{24}O_{14}^{+}$ $[M+H]^{+}$: 1545.7616; found: 1545.107. HPLC: t_R=12.67 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–20 min).



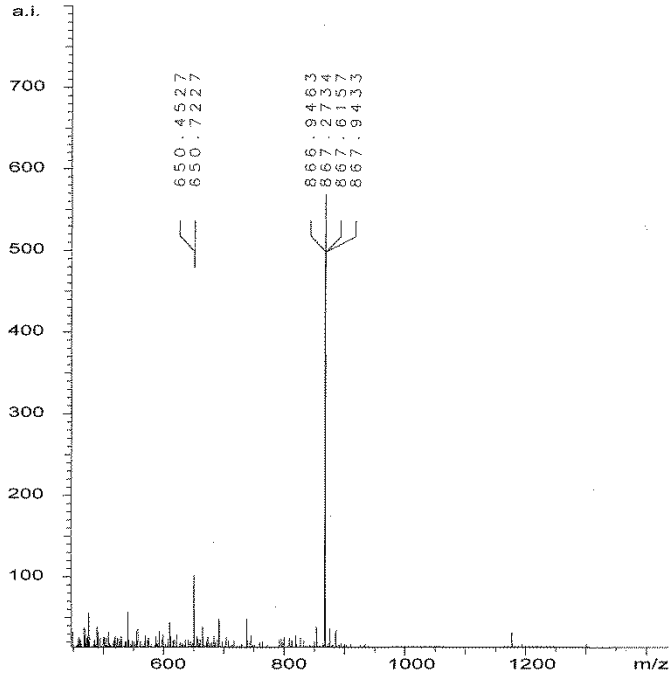
Ada2 was obtained as a white powder. ESI-TOF MS: *m/z* calcd for $C_{80}H_{110}N_{25}O_{17}^{3+}$ $[M+3H]^{3+}$: 564.2837; found: 564.2218, 564.5593, 564.9067; $C_{80}H_{109}N_{25}O_{17}^{2+}$ $[M+2H]^{2+}$: 845.9216; found: 845.8101, 846.3149, 846.8141. HPLC: t_R=12.78 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–20 min).



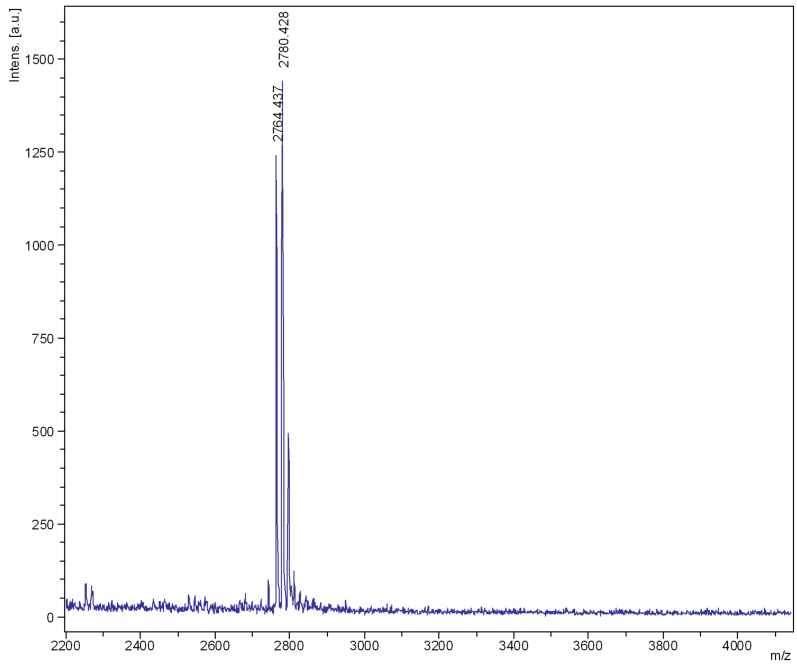
Ada3 was obtained as a white powder. MALDI-TOF MS: *m/z* calcd for $C_{80}H_{105}N_{26}NaO_{16}$ $[M+Na]^+$: 1709.876; found: 1709.975. HPLC: t_R=20.325 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–40 min).



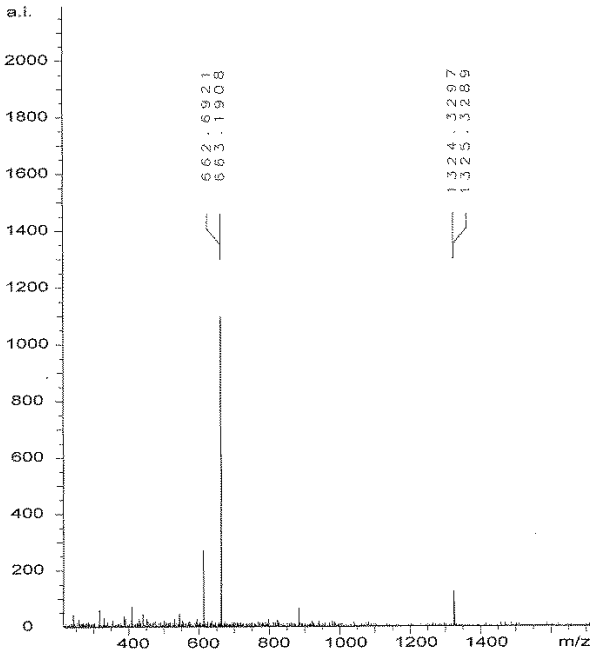
CyD1 was obtained as a white powder. ESI-TOF MS: *m/z* calcd for $C_{110}H_{163}N_{25}O_{48}^{4+}$ $[M+4H]^{4+}$: 650.7779; found: 650.4527, 650.7227; $C_{110}H_{162}N_{25}O_{48}^{3+}$ $[M+3H]^{3+}$: 867.3679; found: 866.9463, 867.2734, 867.6157. HPLC: t_R=9.50 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–20 min).



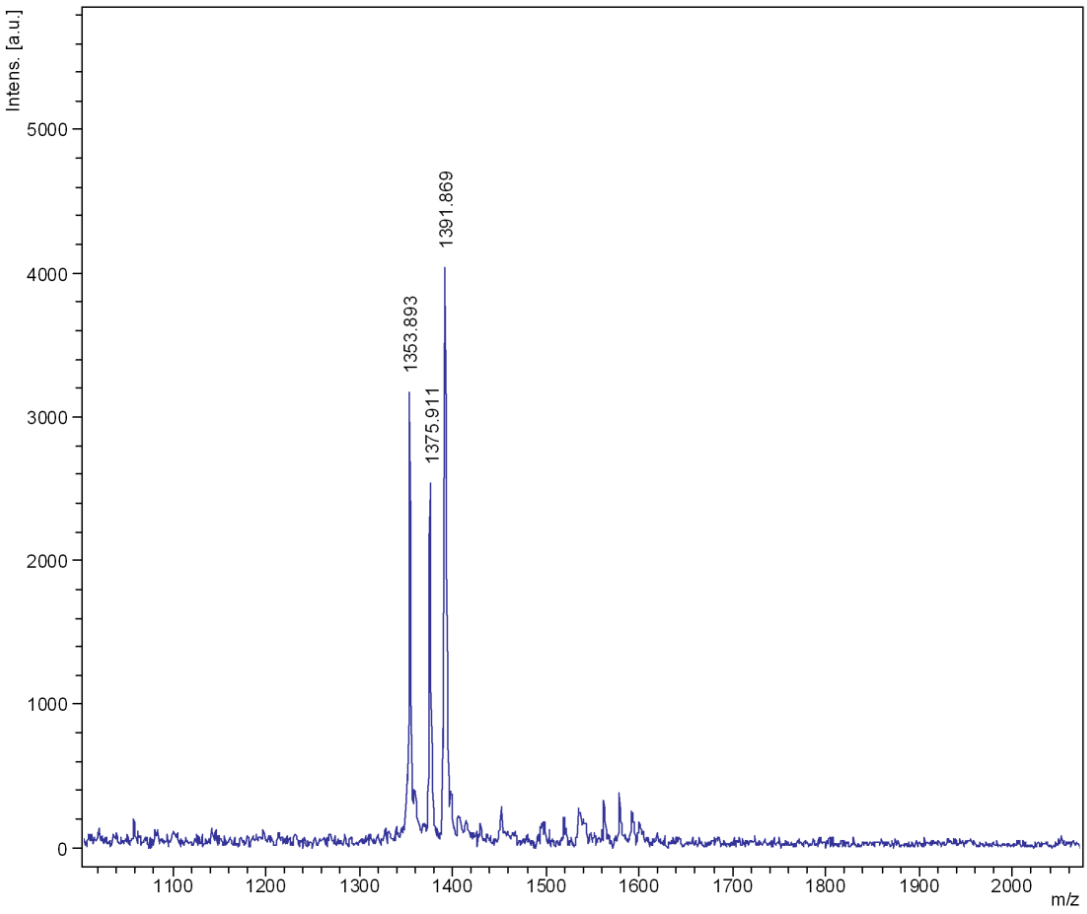
Cyd2 was obtained as a white powder. MALDI-TOF MS: *m/z* calcd for $C_{116}H_{168}N_{27}NaO_{50}^{+}$ $[M+Na]^{+}$: 2763.749; found: 2764.437. HPLC: t_R=14.775 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–40 min).



PIP1 was obtained as a white powder. ESI-TOF MS: *m/z* calcd for $C_{61}H_{78}N_{23}O_{12}^{+}$ $[M+H]^{+}$: 1324.6195; found: 1324.3297, 1325.3289; $C_{61}H_{78}N_{23}O_{12}^{2+}$ $[M+2H]^{2+}$ found: 662.6921, 663.1908. HPLC: t_R=11.675 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–20 min). (Mass data was attached in the bottom)



PIP2 was obtained as a white powder. MALDI-TOF MS: *m/z* calcd for $C_{62}H_{80}N_{23}O_{13}^{+}$ $[M+H]^{+}$: 1354.6300; found: 1353.893; $C_{61}H_{78}N_{23}O_{12}^{2+}$ $[M+Na]^{+}$ found: 1375.911. HPLC: t_R=11.208 min (0.1□% TFA/MeCN, linear gradient 0–100□%, 0–20 min). (Mass data was attached in the bottom)



References:

(1) Tang, W.; Ng, S.-C. *Nat. Protocols* **2007**, *2*, 3195-3200.

(2) Liu, Y.-Y.; Fan, X.-D.; Gao, L. *Macromol. Biosci.* **2003**, *3*, 715-719.

(3) Yu, Z.; Taniguchi, J.; Wei, Y.; Pandian, G. N.; Hashiya, K.; Bando, T.; Sugiyama, H. *Eur. J. Med. Chem.* **2017**, *138*, 320-327.

(4) Guo, C.; Kawamoto, Y.; Asamitsu, S.; Sawatani, Y.; Hashiya, K.; Bando, T.; Sugiyama, H. *Bioorg. Med. Chem.* **2015**, *23*, 855-860.

(5) Heddi, B.; Cheong, V. V.; Martadinata, H.; Phan, A. T. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 9608-9613.

(6) Asamitsu, S.; Li, Y.; Bando, T.; Sugiyama, H. *Chembiochem* **2016**, *17*, 1317-1322.

(7) Wilson, V. S.; Bobseine, K.; Gray, J. L. E. *Toxicological Sciences* **2004**, *81*, 69-77.

(8) Chou, H.-M.; Chao, H.-R.; Lin, C.; Chiang, P.-C.; Wang, G.-S.; Tsou, T.-C. *Toxicological & Environmental Chemistry* **2016**, *98*, 376-384.

(9) Summerton, J. E. *Annals of the New York Academy of Sciences* **2005**, *1058*, 62-75.

(10) Schagat, T.; Paguio, A.; Kopish, K. *Normalizing genetic reporter assays: approaches and considerations for increasing consistency and statistical significance*, 2007; Vol. 17.

(11) Guo, C.; Asamitsu, S.; Kashiwazaki, G.; Sato, S.; Bando, T.; Sugiyama, H. *ChemBioChem* **2017**, *18*, 166-170.

(12) Chenoweth, D. M.; Dervan, P. B. *J. Am. Chem. Soc.* **2010**, *132*, 14521-14529.

(13) van den Berg, B.; Prathyusha Bhamidimarri, S.; Dahyabhai Prajapati, J.; Kleinekathöfer, U.; Winterhalter, M. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, E2991-E2999.

(14) Dunn, A. R.; Hays, A. M.; Goodin, D. B.; Stout, C. D.; Chiu, R.; Winkler, J. R.; Gray, H. B. *J. Am. Chem. Soc.* **2002**, *124*, 10254-10255.

(15) Laughlin-Toth, S.; Carter, E. K.; Ivanov, I.; Wilson, W. D. *Nucleic Acids Res.* **2017**, *45*, 1297-1306.